View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Institutional Knowledge at Singapore Management University

Singapore Management University
Institutional Knowledge at Singapore Management University

Research Collection Lee Kong Chian School Of Business

Lee Kong Chian School of Business

9-2008

Pareto-optimal predictor composite formation: A complementary approach to alleviating the selection quality/adverse impact dilemma

Paul R. SACKETT
University of Minnesota - Twin Cities

Wilfried DE CORTE
Ghent University

Filip LIEVENS
Singapore Management University, filiplievens@smu.edu.sg
DOI: https://doi.org/10.1111/j.1468-2389.2008.00426.x

Follow this and additional works at: https://ink.library.smu.edu.sg/lkcsb_research

Part of the Human Resources Management Commons, and the Organizational Behavior and Theory Commons

Citation

SACKETT, Paul R.; DE CORTE, Wilfried; and LIEVENS, Filip. Pareto-optimal predictor composite formation: A complementary approach to alleviating the selection quality/adverse impact dilemma. (2008). *International Journal of Selection and Assessment*. 16, (3), 206-209. Research Collection Lee Kong Chian School Of Business.
Available at: https://ink.library.smu.edu.sg/lkcsb_research/5647

This Journal Article is brought to you for free and open access by the Lee Kong Chian School of Business at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection Lee Kong Chian School Of Business by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email libIR@smu.edu.sg.

Published in International Journal of Selection and Assessment, 2008 September, Volume 16, Issue 3, Pages 206-209

Pareto-Optimal Predictor Composite Formation: A complementary approach to alleviating the selection quality/ adverse impact dilemma

Paul R. Sackett*, Wilfried De Corte and Filip Lievens*****

*Department of Psychology, University of Minnesota, Twin Cities Campus, Minneapolis, MN 55455, USA
**Department of Data Analysis, Ghent University, H. Dunantlaan 1, 9000 Ghent, Belgium. wilfried.decorte@UGent.be
***Department of Personnel Management and Work and Organizational Psychology, Ghent University, Ghent, Belgium

In our rejoinder to the comments of Kehoe (this issue) and Potosky, Bobko and Roth (this issue) we emphasize that our proposal on Pareto-optimal predictor composite formation is a complementary and not a competitive alternative for reducing the tension between selection quality and adverse impact. Our work addresses the decisions to be made once one has decided to use a predictor composite. We also further clarify the basic features of Pareto-optimal tradeoffs and Pareto-optimal composites within the context of personnel selection. In particular, we indicate that Pareto-optimal tradeoffs between validity and adverse impact emerge because these goals are different and not because of any dualism between them.

We welcome the opportunity for an exchange of views regarding our paper, discussing the use of Pareto-optimization methods to shed light on the question of the degree to which changes in the weights assigned to a chosen set of predictors jointly affect the outcomes of composite validity and adverse impact (De Corte, Lievens, & Sackett, 2008). We are pleased that Kehoe (2008) and Potosky, Bobko, and Roth (2008) accepted the editor's invitation to comment on our work, as we believe the exchange broadens the reader's insight into and understanding of our work. We note at the outset that we do not see significant points of disagreement with Kehoe or with Potosky *et al*. Kehoe brings a useful applied perspective to the issue, raising several additional issues that expand on our work. Potosky *et al*.'s remarks serve to identify potential misinterpretations of our work and thus provide us with the opportunity to clarify the scope and intent of our work.

Kehoe first notes that our approach, which involves an explicit examination of the degree to which predictor weighting schemes that result in a lowering of validity from the degree maximally attainable with that predictor set can achieve increases in the adverse impact ratio, may not be appealing to some organizations. He acknowledges that some organizations may wish to frame the issue in terms of this explicit framework; our work is aimed at those who make this

choice. It is very useful to also appreciate Kehoe's description of why other organizations may prefer not to call this attention to the issue and why they may prefer a wide variety of other approaches to seeking a more diverse work force, such as increased recruiting efforts aimed at increasing the quality of the minority applicant pool. Again, we do not offer our approach as the sole approach, or as the best approach, to pursuing diversity; rather, we simply observe that the choice of predictor weights is one decision that can affect validity and adverse impact, and we offer a systematic approach to considering the effects of weighting choices.

Kehoe's second concern is whether the validity figures presented in our paper fully and correctly represent the value to the organization. He suggests that it is best to view our paper as a heuristic device. We agree, as our paper uses one meta-analytically compiled data set using a set of four predictors to illustrate our approach. In various applied settings, differing predictors will be used and different meta-analytic mean values will be appropriate (e.g., the validity of cognitive ability (CA) tests varies by job complexity). That said, Kehoe does raise two provocative and important points that merit attention in future work. The first is to question the meta-analytic mean correlations we use for CA and conscientiousness. The values used reflect the 'state of the science' in our published literature, and thus Kehoe's challenge is to the field as a whole, not simply to our paper's use of these values. The meta-analytic values are drawn from studies correlating CA and conscientiousness with measures of overall job performance. If we follow the argument correctly, Kehoe appears to suggest that the criterion measures used in studies of CA will overemphasize task performance, and measures used in studies of conscientiousness will overemphasize non-task performance (e.g., citizenship, interpersonal or 'soft' aspects of the job). It is certainly true that one can only make meaningful comparisons of validity when criteria are comparable, and thus Kehoe's challenge is an important one. Note that Kehoe offers this as an impressionistic observation, not as a documented fact; clearly, research on this issue is needed.

Kehoe also raises the concern that the predictive validity of CA with regard to the criterion of overall job performance may not fully capture the value to the organization of screening on cognitive ability. He offers the observation that ability is also related to health behaviors, and thus screening on CA can have the incidental benefit of a healthier workforce. Kehoe is certainly correct here: if a given predictor has value to an organization through a link to multiple outcomes, all of those outcomes need to be considered in any consideration of tradeoffs. We note that Kehoe's ideas here can contribute to broader thinking in the field. Note, for example, that Kehoe's broader concerns are also ignored by traditional regression-based approaches to predictor selection and weighting, as the decision to use and weight a predictor from a trial battery is based solely on its link to the measured job performance criterion. Yet both the regression-based and our Pareto-optimal weighting approach can account for the multidimensional nature of valued criterion behavior when data are available on the predictor validities and effect sizes with regard to the different criterion aspects as well as on the intercorrelation between the criterion dimensions. Thus, De Corte, Lievens, and Sackett (2007) study Pareto-optimal composites where the valued criterion behavior is an aggregate of both task and contextual performance aspects. Other criterion aspects such as health and safety behaviors mentioned by Kehoe can be integrated as well, provided that relevant data become available. Here again, additional research leading to adequate meta-analytic estimates is more than welcome.

Kehoe's third point is that even if organizations are willing to consider validity–adverse impact tradeoffs, they may wish to consider a broader array of approaches other than simply the weighting of predictors. Kehoe notes the use of cutoff scores, rather than the compensatory approach implied by selecting on a weighted predictor composite, and the use of banding as additional alternatives to consider. Again, we support a broad investigation of alternatives; our work addresses the decisions to be made once one has decided to use a predictor composite, while Kehoe addresses alternatives to using composites.

Potosky *et al.* (2008) express a series of concerns about our work. One is that they interpret our work as indicating that those who use regression weights (i.e., thus maximizing validity) do not care about adverse impact or would not value a reduction in adverse impact. We note that we do not intend any attributions about personal motives or values; rather, the course of action chosen does not reflect explicit concern for adverse impact. As explained more fully hereafter, the regression-based composite is a particular element of the set of all Pareto-optimal composites one obtains when both validity and adverse impact are of concern.

A second concern is that the word 'optimal' has a common connotation, which is different from its use in the context of the framework of Pareto-optimality. We agree that Pareto-optimality is a very different concept, and our experiences in presenting our work suggest that it is a concept unfamiliar to many in the selection field. Thus, it is indeed critical that our work is clearly presented and clearly understood. To correct possible misinterpretations, we recapitulate and further clarify the basic features of Pareto-optimal trade-offs in the context of personnel selection. We start by noting that the quest for predictor composites that aim for maximum validity and maximum diversity represents a

particular example of multi-objective optimization. In general, multi-objective optimization seeks solutions that jointly optimize several outcomes, where the value of each outcome depends on the particular values of certain relevant design or decision variables. Thus, in the present context, the design variables correspond to the weights with which the available predictors are combined to composites, and the composite validity and the expected adverse impact ratio when implementing the composite represent the two outcomes. Unfortunately, for this as well as for virtually any other instance of multi-objective optimization, the optimum value of the objectives is not obtained for the same set of values for the design variables such that no solution exists that simultaneously optimizes the objectives. Instead, many different, so-called Pareto-optimal, solutions are obtained. Each Pareto-optimal solution corresponds to a set of values for the design variables, say the set *S* of design values, such that no other set of design values can improve at least one of the objective values associated with the set *S* of design values and, at the same time, lead to equal or better values for the other objectives. For example, while there may be many composites that would produce the same level of validity, the Pareto-optimal solution at that level of validity is the one that produces the most favorable adverse impact ratio. Thus, seeking predictor weighting systems that optimize both validity and adverse impact ratio leads to many different Pareto-optimal weighting systems, and each such system corresponds to a particular pair of values for the validity and the adverse impact ratio objective.

Conforming to the general literature on multi-objective optimization and decision-making, in our paper we used the term 'Pareto-optimal tradeoff' to refer to the set of objective values that corresponds to a Pareto-optimal weighting system and called the resulting composite Pareto-optimal as well. Also, from the total of Pareto-optimal composites, one corresponds to the maximum validity composite (i.e., the regression-weighted composite), another corresponds to the minimum impact composite (i.e., the composite with the highest possible adverse impact ratio) and all others represent 'balanced' tradeoffs between validity and adverse impact ratio. If one is interested in only one of the objectives, then either the regression composite or the minimal impact composite is the optimal solution, depending on whether the valued objective is validity or adverse impact ratio maximization. When both the objectives are judged important, then all Pareto-optimal composites are of potential interest and there is no automatic, objective procedure to choose one of the composites as *the* optimal one. Also, in that case, all other composites – that is, all composites that do not correspond to a Pareto-optimal weighting system – can be dismissed. Hence, when addressing the issue of predictor composite formation in a context where both validity and diversity are valued, one ends up with a clear division between composites that merit further attention (i.e., all and only the Pareto-optimal composites) and all others that do not.

Observe that obtaining Pareto-optimal trade-offs between validity and adverse impact ratio has nothing to do with modeling 'adverse impact and validity as if they are inversely related' as suggested by Potosky *et al.* (2008) in their third comment. Adverse impact and validity are simply two different objectives. Whether they are (inversely) linked, either conceptually or empirically, is essentially of no relevance. All that matters is that no single weight system maximizes both validity and the adverse impact ratio. Hence, our proposal on Pareto-optimal tradeoffs does not reflect a position where validity maximization is placed at one end of a goal continuum and adverse impact at the other. We do not set up a dualism between these two objectives, but simply recognize that these are different goals. In fact, even with predictors that show a proportional relation between validity and effect size, there will typically be a difference between the weights (and the resulting composite) that maximize validity and the weight system that corresponds to maximizing the adverse impact ratio, resulting once again in a set of Pareto-optimal tradeoffs.

Potosky *et al.* (2008) also express the concern that we offer little guidance as to how to make what both they and we acknowledge as a value-based decision. We note that the fact that it is a value-based decision makes it one where it is not our place to tell someone which decision to make. But we believe that our approach is useful in informing the value decision, as it helps make clear what can and cannot be achieved by changing predictor weights. In some cases, the finding may be that no change in weights can meaningfully improve adverse impact without a large loss in validity; in other cases, the finding may be that adverse impact might be improved substantially with a small loss in validity. Of course, what is 'small' to one may be 'large' to another, which is what makes this a value judgment. But we believe that a clear statement of what can and cannot be achieved by alternative weighting schemes can be a useful input to this decision.

We agree with Kehoe (p. 8) and Potosky *et al.* (point 4) that choosing between the different Pareto-optimal tradeoffs may often be made more transparent by converting the validity value of the Pareto-optimal trade-offs into another more meaningful value such as the utility or the expected mean performance of the selected applicants. These conversions require additional data, however, such as, for example estimates on the recruiting and predictor costs as well as on the monetary value of criterion performance levels in case

of the utility metric. When feasible, a retranslation in terms of these other metrics may offer a clearer understanding of the loss in selection quality embraced by any particular balanced Pareto-optimal trade-off. Useful translations that do not require additional data are also possible. Kehoe (2008) offers such a translation by suggesting to characterize Pareto-optimal composites that include a CA predictor in terms of the average CA score of the composite selected applicants.

Potosky *et al*. (2008) express a final concern that our method has not 'solved' the adverse impact problem and that its application still produces 'sobering' results. We certainly agree that our proposal does not solve the adverse impact problem; in no way did we ever suggest that we had. As noted repeatedly throughout this commentary, we offer an approach to efficiently evaluating what can and cannot be accomplished by alternate predictor weighting schemes, which we view as a modest, but useful, contribution to selection system design and implementation. Also, whether the results of applying the method are sobering or not very much depends on the framing of these results. When looking from the Potosky *et al*.'s point of view, which focuses on the additional number of minority hires in a selective screening (i.e., a 20 percent overall selection rate) from a fixed-sized applicant group that is predominantly composed of majority candidates (i.e., a .881 representation of majority applicants), the reported results are indeed far from impressive. However, we favor reporting the increase in minority hires by means of the % minority representation index that we also use in our paper because this index translates the expected effect of using a balanced Pareto-optimal composite over a series of applications. Also, consider the following analogy. Suppose that we want to assess the value of a new drug to treat cancer. Would we then say that the new drug produces sobering results when its administration results in a 41% increase of survival as compared with the previous drug? Yet, this is the actual increase for the example used by Potosky *et al*. (p. 10) to illustrate that the results of choosing a balanced Pareto-optimal composite instead of the regression-based composite are sobering.

All in all, we believe that the stimulating comments by Kehoe (2008) and Potosky *et al*. (2008) offered us the opportunity to better articulate the nature and the extent of our proposal on Pareto-optimal predictor composite formation. When perceived as a complementary, instead of a competitive proposal for reducing the tension between selection quality and adverse impact, we think that considering 'balanced' Pareto-optimal composites besides the regression-based composite offers a worthwhile effort.

References

De Corte, W., Lievens, F. and Sackett, P. (2007) Combining Predictors to Achieve Optimal Trade-Offs Between Selection Quality and Adverse Impact. *Journal of Applied Psychology*, **92**, 1380–1393.

De Corte, W., Lievens, F. and Sackett, P. (2008) Validity and Adverse Impact Potential of Predictor Composite Formation. *International Journal of Selection and Assessment*, **16**, 183–194.

Kehoe, J.F. (2008) Commentary on Pareto-Optimality as a Rationale for Adverse Impact Reduction: What would organizations do? *International Journal of Selection and Assessment*, **16**, 195–200.

Potosky, D., Bobko, P. and Roth, P.L. (2008) Some Comments on Pareto Thinking, Test Validity, and Adverse Impact: When "and" is optimal and "or" is a trade-off. *International Journal of Assessment and Selection*, **16**, 201–205.